

## (1) GENERAL INFORMATION

(ii) TITLE OF THE INVENTION: HUMAN MEMBRANE RECYCLING PROTEINS

(iii) NUMBER OF SEQUENCES: 5

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Incyte Pharmaceuticals, Inc.

(B) STREET: 3174 Porter Dr.

(C) CITY: Palo Alto

(D) STATE: CA

(E) COUNTRY: USA

(F) ZIP: 94304

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette

(B) COMPUTER: IBM Compatible

(C) OPERATING SYSTEM: DOS

(D) SOFTWARE: FastSEQ for Windows Version 2.0

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: To Be Assigned

(B) FILING DATE: Filed Herewith

(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 09/360,125

(B) FILING DATE: July 23, 1999

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 09/004,502

(B) FILING DATE: January 8, 1998

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Billings, Lucy J.

(B) REGISTRATION NUMBER: 36,749

(C) REFERENCE/DOCKET NUMBER: PF-0456 US

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 650-855-0555

(B) TELEFAX: 650-845-4166

## (2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 347 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

(A) LIBRARY: TONGTUT01

(B) CLONE: 980615

(xi) SEQUENCE DESCRIPTION: SEO ID NO:1:

59

20 25 30  
 Arg Gln Tyr Ala Thr Leu Asp Val Tyr Asn Pro Phe Glu Thr Arg Glu  
 35 40 45  
 Pro Pro Pro Ala Tyr Glu Pro Pro Ala Pro Ala Pro Leu Pro Pro Pro  
 50 55 60  
 Ser Ala Pro Ser Leu Gln Pro Ser Arg Lys Leu Ser Pro Thr Glu Pro  
 65 70 75 80  
 Lys Asn Tyr Gly Ser Tyr Ser Thr Gln Ala Ser Ala Ala Ala Thr  
 85 90 95  
 Ala Glu Leu Leu Lys Lys Gln Glu Glu Leu Asn Arg Lys Ala Glu Glu  
 100 105 110  
 Leu Asp Arg Arg Glu Arg Glu Leu Gln His Ala Ala Leu Gly Gly Thr  
 115 120 125  
 Ala Thr Arg Gln Asn Asn Trp Pro Pro Leu Pro Ser Phe Cys Pro Val  
 130 135 140  
 Gln Pro Cys Phe Phe Gln Asp Ile Ser Met Glu Ile Pro Gln Glu Phe  
 145 150 155 160  
 Gln Lys Thr Val Ser Thr Met Tyr Tyr Leu Trp Met Cys Ser Thr Leu  
 165 170 175  
 Ala Leu Leu Leu Asn Phe Leu Ala Cys Leu Ala Ser Phe Cys Val Glu  
 180 185 190  
 Thr Asn Asn Gly Ala Gly Phe Gly Leu Ser Ile Leu Trp Val Leu Leu  
 195 200 205  
 Phe Thr Pro Cys Ser Phe Val Cys Trp Tyr Arg Pro Met Tyr Lys Ala  
 210 215 220  
 Phe Arg Ser Asp Ser Ser Phe Asn Phe Phe Val Phe Phe Phe Ile Phe  
 225 230 235 240  
 Phe Val Gln Asp Val Leu Phe Val Leu Gln Ala Ile Gly Ile Pro Gly  
 245 250 255  
 Trp Gly Phe Ser Gly Trp Ile Ser Ala Leu Val Val Pro Lys Gly Asn  
 260 265 270  
 Thr Ala Val Ser Val Leu Met Leu Leu Val Ala Leu Leu Phe Thr Gly  
 275 280 285  
 Ile Ala Val Leu Gly Ile Val Met Leu Lys Arg Ile His Ser Leu Tyr  
 290 295 300  
 Arg Arg Thr Gly Ala Ser Phe Gln Lys Ala Gln Gln Glu Phe Ala Ala  
 305 310 315 320  
 Gly Val Phe Ser Asn Pro Ala Val Arg Thr Ala Ala Ala Asn Ala Ala  
 325 330 335  
 Ala Gly Ala Ala Glu Asn Ala Phe Arg Ala Pro  
 340 345

## (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1521 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (vii) IMMEDIATE SOURCE:

- (A) LIBRARY: TONGTUT01  
 (B) CLONE: 980615

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

NGACGCAGGC	GCAACCCACG	GCTGCTGCGG	GGATCCTTGT	GGCCCTTCCG	GTGCATGGAA	60
CCAATCCGTG	CACAGAGAAG	CGGGGCGAAC	TGAGGCGAGT	GAAGTGGACT	CTGAGGGGCTA	120
CCGCTACCGC	CACTGCTGCG	GCAGGGGCGT	GGAGGGCAGA	GGGCCGCGGA	GGCCGCGAGTT	180
GCAAACATGG	CTCAGAGCAG	AGACGGCGGA	AACCCGTTTCG	CCGAGCCCAG	CGAGCTTGAC	240
AACCCCTTTC	AGGACCCAGC	TGTGATCCAG	CACCGACCCA	GCCGCGCAGTA	TGCCACGCTT	300
GACGTCTACA	ACCCTTTTGA	GACCCGGGAG	CCACCACCAG	CCTATGAGCC	TCCAGCCCCCT	360
GCCCCATTGC	CTCCACCCTC	AGCTCCCTCC	TTGCAGCCCT	CGAGAAAGCT	CAGCCCCACA	420
GAACCTAAGA	ACTATGGCTC	ATACAGCACT	CAGGCCTCAG	CTGCAGCAGC	CACAGCTGAG	480
CTGCTGAAGA	AACAGGAGGA	GCTCAACCGG	AAGGCAGAGG	AGTTGGACCG	AAGGGAGCGA	540
GAGCTGCAGC	ATGCTGCCCT	GGGGGGCACA	GCTACTCGAC	AGAACAATTG	GCCCCCTCTA	600
CCTTCTTTTT	GTCCAGTTCA	GCCCTGCTTT	TTCCAGGACA	TCTCCATGGA	GATCCCCCAA	660

GAATTTTCAGA	AGACTGTATC	CACCATGTAC	TACCTCTGGA	TGTGCAGCAC	GCTGGCTCTT	720
CTCCTGAACT	TCCTCGCCTG	CCTGGCCAGC	TTCTGTGTGG	AAACCAACAA	TGGCGCAGGC	780
TTTGGGCTTT	CTATCCTCTG	GGTCCCTCTT	TTCACTCCCT	CTGCTGGTAC		840
CGCCCCATGT	ATAAGGCTTT	CCGGAGTGAC	AGTTCATTCA	ATTTCTTCGT	TTTCTTCTTC	900
ATTTTCTTCG	TCCAGGATGT	GCTCTTTGTC	CTCCAGGCCA	TTGGTATCCC	AGGTTGGGGA	960
TTCAGTGGCT	GGATCTCTGC	TCTGGTGGTG	CCGAAGGGCA	ACACAGCAGT	ATCCGTGCTC	1020
ATGCTGCTGG	TCGCCCTGCT	CTTCACTGGC	ATTGCTGTGC	TAGGAATTGT	CATGCTGAAA	1080
CGGATCCACT	CCTTATACCG	CCGCACAGGT	GCCAGCTTTC	AGAAGGCCCA	GCAAGAATTT	1140
GCTGCTGGTG	TCTTCTCCAA	CCCTGCGGTG	CGAACCGCAG	CTGCCAATGC	AGCCGCTGGG	1200
GCTGCTGAAA	ATGCCCTCCG	GGCCCCGTGA	CCCCTGACTG	GGATGCCCTG	GCCCTGCTAC	1260
TTGAGGGAGC	TGACTTAGCT	CCCGTCCCTA	AGGTCTCTGG	GACTTGGAGA	GACATCACTA	1320
ACTGATGGCT	CCTCCGTAGT	GCTCCCAATC	CTATGGCCAT	GACTGCTGAA	CCTGACAGGC	1380
GTGTGGGGAG	TTCACGTGTGA	CCTAGTCCCC	CCATCAGGCC	ACACTGCTGC	CACCTCTCAC	1440
ACGCCCCAAC	CCAGCTTCCC	TCTGCTGTGC	CACGGCTGTT	GCTTCGGTTA	TTTAAATAAA	1500
AAGAAAGTGG	AACTGGAAGT	G				1521

## (2) INFORMATION FOR SEQ ID NO:3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 329 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (vii) IMMEDIATE SOURCE:

- (A) LIBRARY: BRSTNOT01  
 (B) CLONE: 412453

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met	Ser	Ala	Phe	Asp	Thr	Asn	Pro	Phe	Ala	Asp	Pro	Val	Asp	Val	Asn
1				5					10					15	
Pro	Phe	Gln	Asp	Pro	Ser	Val	Thr	Gln	Leu	Thr	Asn	Ala	Pro	Gln	Gly
			20					25					30		
Gly	Leu	Ala	Glu	Phe	Asn	Pro	Phe	Ser	Glu	Thr	Asn	Ala	Ala	Thr	Thr
		35					40					45			
Val	Pro	Val	Thr	Gln	Leu	Pro	Gly	Ser	Ser	Gln	Pro	Ala	Val	Leu	Gln
	50					55					60				
Pro	Ser	Val	Glu	Pro	Thr	Gln	Pro	Thr	Pro	Gln	Ala	Val	Val	Ser	Ala
65					70					75				80	
Ala	Gln	Ala	Gly	Leu	Leu	Arg	Gln	Gln	Glu	Glu	Leu	Asp	Arg	Lys	Ala
				85					90					95	
Ala	Glu	Leu	Glu	Arg	Lys	Glu	Arg	Glu	Leu	Gln	Asn	Thr	Val	Ala	Asn
		100						105					110		
Leu	His	Val	Arg	Gln	Asn	Asn	Trp	Pro	Pro	Leu	Pro	Ser	Trp	Cys	Pro
	115						120					125			
Val	Lys	Pro	Cys	Phe	Tyr	Gln	Asp	Phe	Ser	Thr	Glu	Ile	Pro	Ala	Asp
	130					135					140				
Tyr	Gln	Arg	Ile	Cys	Lys	Met	Leu	Tyr	Tyr	Leu	Trp	Met	Leu	His	Ser
145					150					155				160	
Val	Thr	Leu	Phe	Leu	Asn	Leu	Leu	Ala	Cys	Leu	Ala	Trp	Phe	Ser	Gly
				165					170					175	
Asn	Ser	Ser	Lys	Gly	Val	Asp	Phe	Gly	Leu	Ser	Ile	Leu	Trp	Phe	Leu
			180					185					190		
Ile	Phe	Thr	Pro	Cys	Ala	Phe	Leu	Cys	Trp	Tyr	Arg	Pro	Ile	Tyr	Lys
		195					200					205			
Ala	Phe	Arg	Ser	Asp	Asn	Ser	Phe	Ser	Phe	Phe	Val	Phe	Phe	Phe	Val
	210					215					220				
Phe	Phe	Cys	Gln	Ile	Gly	Ile	Tyr	Ile	Ile	Gln	Leu	Val	Gly	Ile	Pro
225					230					235				240	
Gly	Leu	Gly	Asp	Ser	Gly	Trp	Ile	Ala	Ala	Leu	Ser	Thr	Leu	Asp	Asn
				245					250					255	
His	Ser	Leu	Ala	Ile	Ser	Val	Ile	Met	Met	Val	Val	Ala	Gly	Phe	Phe
		260					265						270		
Thr	Leu	Cys	Ala	Val	Leu	Ser	Val	Phe	Leu	Leu	Gln	Arg	Val	His	Ser
	275						280					285			
Leu	Tyr	Arg	Arg	Thr	Gly	Ala	Ser	Phe	Gln	Gln	Ala	Gln	Glu	Glu	Phe

290 295 300  
 Ser Gln Gly Ile Phe Ser Ser Arg Thr Phe His Arg Ala Ala Ser Ser  
 305 310 315 320  
 Ala Ala Gln Gly Ala Phe Gln Gly Asn  
 325

## (2) INFORMATION FOR SEQ ID NO:4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2434 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (vii) IMMEDIATE SOURCE:

- (A) LIBRARY: BRSTNOT01  
 (B) CLONE: 412453

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

NCCGGAAGTG GAGGGTCTAC ACGAAGCGCC GCTGGGTCTG GGTGCCCCGGA GGCAGCAGCG 60  
 TTCGCGGAGT TCGCCCGCTG GCCCCCGATC ACCATGTCTG CTTTCGACAC CAACCCCTTC 120  
 GCGGACCCAG TGGATGTAAA CCCCTTCCAG GATCCCTCTG TGACCCAGCT GACCAACGCC 180  
 CCGCAGGGCG GCCTGGCGGA ATTCAACCCC TTCTCAGAGA CAAATGCAGC GACAACAGTT 240  
 CCTGTCACCC AACTCCCTGG GTCTTCACAG CCAGCGGTTT TCCAGCCATC AGTGGAAACCA 300  
 ACCCAGCCGA CCCCCCAGGC CGTGGTGTCT GCAGCCAGG CAGGCCTGCT CCGGCAGCAG 360  
 GAAGAACTGG ACAGGAAAGC TGCCGAGCTG GAACGCAAGG AGCGGGAGCT GCAGAACT 420  
 GTAGCCAACT TGCATGTGAG ACAGAAACAC TGGCCCCCTG TGCCCTCGTG GTGCCCTGTG 480  
 AAGCCCTGCT TCTATCAGGA TTTCTCCACA GAGATCCCTG CCGACTACCA GCGGATATGC 540  
 AAGATGCTCT ACTATCTGTG GATGTTGCAT TCAGTGACTC TGTTTCTGAA CCTGCTTGCC 600  
 TGCCTGGCCT GGTTCCTCGG CAACAGCTCC AAGGGAGTGG ACTTTGGCCT CTCCATCCTG 660  
 TGGTTTCTGA TCTTCACTCC CTGTGCCTTC CTTTGTGTTG ACCGACCCAT CTATAAGGCC 720  
 TTTAGGTCCG ACAACTCTTT CAGCTTCTTT GTGTTCTTCT TTGTATTTTT TTGTCAAATA 780  
 GGGATCTACA TCATCCAGTT GGTTCGCATC CCTGGCCTGG GGGACAGCGG TTGGATTGCA 840  
 GCCCTGTCTA CACTGGATAA TCATTCCCTG GCCATATCAG TCATCATGAT GGTGGTGGCT 900  
 GGCTTCTTCA CCCTCTGTGC CGTGCTCTCA GTCTTCCTCC TGCAGCGGGT GCACTCCCTC 960  
 TACCGACGGA CAGGGGCCAG CTTCCAGCAG GCCCAGGAGG AGTTTCCCA GGGCATCTTC 1020  
 AGCAGCAGAA CCTTCCACAG AGCTGCTTCA TCTGCTGCCC AAGGAGCCTT CCAGGGGAAT 1080  
 TAGTCCTCCT CTCTTCTCTC CCCCTCAGCC TTTCTCTCGC CTGCCTTCTG AGCTGCACTT 1140  
 TCCGTGGGTG CCTTATGTGG TGGTGGTTGT GCCCAGACA GACCTGGCAG GGTTCCTGCG 1200  
 GTGGCTCTTC CTCTCCCTC AGCGACCAGC TCTCCCTGGA ACGGGAGGGA CAGGGAATTT 1260  
 TTTCCCCCTC TATGTACAAA AAAAAACAAA GCTCTCTTTC CTTCTCTGGT GATGGTTTGG 1320  
 TAGGATTCTT TTGTCTCTGG AAGCAGTGGG ACTGAAGTTC TCTTCGTCCT GTGCACACAC 1380  
 AGACACCCCC ACACAGTTGG GATCACAGGC TGACCTGGGC CCATCCCAGC TGGAGCTTTC 1440  
 TGCCAGGGTC CTGGGCCCTG ACTCCCCCAC CCTGCAGGCC TGGCCTGAAT CTGGCTTCTT 1500  
 AGACACAGCC CAGTCCTTCC TGCCCTGGGT GGGGAATAAG CTCTCACAGG TTCTGGTGGA 1560  
 CAGATCTGTT CCCCAGGTCA CTCCAGTGGT TCCAGGCTT CCAGAGAAGG CTGGTTGCCT 1620  
 CAAGCTCTTC TCTGCCTCAT AAACGGATCC AGAGAAGGCT GGTTCCTTA AGCTCTTCCC 1680  
 TGCCTCGTGT TCCTGAGAAA CGGATTAATA GCCCTTTATC CCCCTGCACC CTCCTGCAGG 1740  
 GGATGGCACT TTGAGCCCTC TGGAGCCCTC CCCTTGCTGA GCCTTACTCT CTTCAGACTT 1800  
 TCTGAATGTA CAGTGCCGTT GGTGGGATT TGGGGACTGG AAGGGACCAA GGACACTGAC 1860  
 CCCAAGCTGT CCTGCCTAGC GTCCAGCGTC TTCTAGGAGG GTGGGGTCTG CCTGTCCTGG 1920  
 TGTGGTTGGT TTGGCCCTGT TTGCTGTGAC TACCCCCCCC CCTCCCCGAA CCGAGGGACG 1980  
 GCTGCCTTTG TCTCTGCCTC AGATGCCACC TGCCCCGCCC ATGCTCCCCA TCAGCAGCAT 2040  
 CCAGACTTTC AGGAAGGGCA GGACCAGCCA GTCCAGAACC GCATCCCTCA GCAGGGACTG 2100  
 ATAAGCCATC TCTCGGAGGG CCCCTAATA CCCAGTGGAG TCTGGTTCAC ACCCTGGGGG 2160  
 GTGTGTCACT GTGATGGGAC ACGTAGGAGT CCACCCTTAA AACCAGCACC CTGTCCCTCG 2220  
 AGGCTGCCGA GTGGGTGTGT GGAAGTGGGT GCCTTCCCAC AAAACTAGCC TCCGGCTCTG 2280  
 GGCCCGAGAC AGCCGAGGC CCCAGCAGT GAATGATACT GGCAGCGGCT GGGGTTTTAT 2340  
 GAACTCCTTT CTGGTATTTT TTCCCCTCTA TGTACAAATG TATATGTTAC GTCTCAATTT 2400  
 TTGTGCTTAA GTAAAAATAA AAACATTTTC AGAC 2434

## (2) INFORMATION FOR SEQ ID NO:5:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 338 amino acids  
 (B) TYPE: amino acid



**THE UNIVERSITY OF CHICAGO**

